

REMARKS

In accordance with the foregoing, the specification and claims 1, 29, 50 and 51 have been amended. Claims 1-33 and 44-62 are pending and under consideration. Claims 30-33, 49 and 58-62 have been allowed.

In item 2 of the Office Action, the Examiner states "[a]pplicant has not amended the specification to remove the new subject matter introduced in the amendment filed 5 June 2003." Applicants are not quite sure of the meaning of item 2. No previous mention of new matter could be found regarding the June 5, 2003 Amendment. However, it appears that the only specification change made on June 5, 2003 is changing " $P(2^{-160})$ " to " $-O(2^{-160})-$ ". To advance prosecution, the original value has been reincorporated.

Applicants are also somewhat confused by item 3 of the Office Action. However, after studying the Examiner's statements, it appears that there is confusion over the use of the words "syntactic" and "syntax." The Manber reference uses the word "syntactic." Applicants used the word "syntax" in the previous Amendment. It appears that the Examiner has checked the definition of these words and believes that "syntactic" or "syntax" could relate to parts of speech. In some circumstances, the Examiner's interpretation would be correct. However, both the Manber reference and the Brin et al. reference are explicitly clear that they are not concerned with parts of speech. Accordingly, the Examiner's interpretation is not correct. It is not appropriate to infer that the references perform a parts of speech analysis simply because the word "syntactic" or "syntax" has been used in connection with the reference. The Examiner is requested to reread the references to eliminate the confusion. The most the references disclose is an analysis of a document based on print layout. A print layout analysis may be concerned with paragraph breaks, which are identified by a new line and an indentation. A print layout analysis may also be concerned with new sentences, which are identified with a period followed by a white space afterwards. However, a print layout analysis is not concerned with parts of speech.

In item 5 of the Office Action, claims 1-20, 44-48 and 50-62 are rejected under 35 USC § 112, second paragraph for indefiniteness. The Examiner asserts that these claims refer to eliminating tokens without any criteria. The claims have been amended to address the rejection. Claim 1 has been amended to clarify that tokens are eliminated based on parts of speech and/or collection statistics. Referring to *Superguide v. DIRECTV*, 69 USPQ2d 1865 (Fed. Cir. 2004), claim 1 does not require the elimination of tokens based on both parts of speech and collection statistics.

Independent claim 29 has been amended to address the indefiniteness rejection by

simply eliminating the objected-to language. For independent claims 50 and 51, these claims previously stated that the document is filtered based on parts of speech. However, to explicitly state how tokens are eliminated, the language has been reordered to state that tokens are eliminated based on parts of speech.

In item 7 of the Office Action, the Examiner sets forth her reasons for allowing claims 30-33, 49 and 58-62. The Examiner states that the prior art does not disclose "a hash algorithm is used for each reordered token in the token stream of the document to generate the hash value for the document." The claims refer to processing retained tokens using a hash algorithm. However, the language mentioned by the Examiner is not contained in the claims. It is believed that the discrepancies between the Examiner's reasons for allowance and the limitations of the independent claims are the result of the Examiner attempting to briefly summarize the limitations of each of the independent claims. It is believed the discrepancies are reflection of the difficulty of this task. It is not believed that the discrepancies are a reflection of an improper understanding on the part of the Examiner. However, if there are any questions regarding the limitations of the independent claims, the Examiner is requested to bring those questions to the attention of the undersigned.

There being no further outstanding objections or rejections, it is submitted that the application is in condition for allowance. An early action to that effect is courteously solicited.

Finally, if there are any formal matters remaining after this response, the Examiner is requested to telephone the undersigned to attend to these matters.

If there are any additional fees associated with filing of this Amendment, please charge the same to our Deposit Account No. 19-3935.

Respectfully submitted,

STAAS & HALSEY LLP

Date: June 20 2005

By: M. J. Henry
Mark J. Henry
Registration No. 36,162

1201 New York Avenue, NW, Suite 700
Washington, D.C. 20005
Telephone: (202) 434-1500
Facsimile: (202) 434-1501